

Uniform Value for Recursive Games with Compact Actions

Xiaoxi Li* and Sylvain Sorin†

April 14, 2016

Abstract

Mertens, Neyman and Rosenberg [6] used the Mertens and Neyman theorem [5] to prove the existence of uniform value for absorbing games with finite state space and compact action sets. We provide an analogous proof for another class of stochastic games, recursive games with finite state space and compact action sets. Moreover, both players have stationary ε -optimal strategies.

Keywords: Stochastic games; Recursive games; Value; Long-term analysis

OR/MS Subject Classification. Primary: games-group decisions/stochastic

1 Introduction

Zero-sum stochastic games were introduced by Shapley [9], and the model $\Gamma = \langle K, I, J, g, q \rangle$ is as follows: K is the finite state space, I (resp. J) is the finite action set for player 1 (resp. player 2), $g : K \times I \times J \rightarrow \mathbb{R}$ is the stage payoff function and $q : K \times I \times J \rightarrow \Delta(K)$ is the probability transition function ($\Delta(K)$ stands for the set of probabilities on K). $k_1 \in K$ is the initial state. At each stage $t \geq 1$, after observing the t -stage history $h_t = (k_1, i_1, j_1, \dots, k_{t-1}, i_{t-1}, j_{t-1}, k_t)$, player 1 chooses an action $i_t \in I$ and player 2 chooses an action $j_t \in J$. This profile (i_t, j_t) induces a current stage payoff $g_t := g(k_t, i_t, j_t)$ and transition probability $q(k_t, i_t, j_t)$ for k_{t+1} the state at the next stage.

We write $\sigma \in \Sigma$ for a behavior strategy for player 1, resp. $\tau \in \mathcal{T}$ for player 2. The λ -discounted value v_λ of the stochastic game Γ (for $\lambda \in (0, 1]$) is defined with respect to the payoff $\gamma_\lambda(k_1, \sigma, \tau) = \mathbb{E}_{\sigma, \tau}^{k_1} [\sum_{t \geq 1} \lambda(1 - \lambda)^{t-1} g(i_t, j_t, k_t)]$. The n -stage value v_n (for $n \geq 1$) is defined analogously by taking the n -stage averaged payoff. Shapley [9] proved that

$$v_\lambda = \Phi(\lambda, v_\lambda), \text{ with} \quad (1.1)$$

$$\forall f \in \mathbb{R}^{|K|}, \quad \Phi(\lambda, f)(k) = \mathbf{val}_{x \in \Delta(I), y \in \Delta(J)} \mathbb{E}_{x, y}^k [\lambda g(i, j, k) + (1 - \lambda)f(k')], \quad (1.2)$$

where $\mathbf{val}_{x \in \Delta(I), y \in \Delta(J)} = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} = \min_{y \in \Delta(J)} \max_{x \in \Delta(I)}$, and that stationary optimal strategies exist for each λ , i.e. depending at each stage t only the current state k_t .

Remark. When the action sets I and J are metric compact, one may assume g and q to be separately continuous on $I \times J$, and this implies the measurability for them (cf. I.1.Ex.7a

*Economics and Management School, Wuhan University, Luojia Hill, 430072 Wuhan, China. Email: xxlee-whu@gmail.com.

†CNRS, IMJ-PRG, UMR 7586, Sorbonne Universités, UPMC Univ Paris 06, Univ Paris Diderot, Sorbonne Paris Cité, Case 247, 4 Place Jussieu, 75252 Paris, France. Email: sylvain.sorin@imj-prg.fr.

in Mertens, Sorin and Zamir [7]). The above description and results extend (cf. VII.1.a in Mertens, Sorin and Zamir [7]): the operator $\Phi(\lambda, \cdot)$ in Equation (1.2) is well defined, v_λ exists and satisfies Equation (1.1), and both players have stationary optimal strategies for each λ .

We are interested in long-run properties of Γ : convergence of v_λ as λ tends to zero or convergence of v_n as n tends to infinity. Moreover, in case of convergence, we ask for the existence of ε -optimal strategies for both players which guarantee the limit value in all games with a sufficiently large horizon n , explicitly:

Definition 1.1 *Let $v \in \mathbb{R}^{|K|}$. Player 1 can guarantee v in Γ if, for any $\varepsilon > 0$ and for every $k_1 \in K$, there exist $N(\varepsilon) \in \mathbb{N}$ and a behavior strategy $\sigma^* \in \Sigma$ for player 1 s.t.: $\forall \tau \in \mathcal{T}$,*

$$(A) \quad \frac{1}{n} \mathbb{E}_{\sigma^*, \tau}^{k_1} \left[\sum_{t=1}^n g_t \right] \geq v(k_1) - \varepsilon, \quad \forall n \geq N(\varepsilon),$$

$$(B) \quad \mathbb{E}_{\sigma^*, \tau}^{k_1} \left[\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n g_t \right] \geq v(k_1) - \varepsilon.$$

A similar definition holds for player 2.

v is the uniform value of Γ if both players can guarantee it.

For stochastic games with finite state space and finite action sets, Bewley and Kohlberg [?] proved the convergence of v_λ as λ tends to zero (and later deduce the convergence of v_n as n tends to infinity), using an algebraic argument. Using the property that the function $\lambda \mapsto v_\lambda$ has bounded variation, that follows from [?], Mertens and Neyman [5] proved the existence of the uniform value. Actually, Mertens and Neyman's theorem [5] is even applicable to a stochastic game Γ with compact action sets under the following form:

Theorem 1.2 (Mertens and Neyman [5]) *Let $\lambda \rightarrow w_\lambda \in \mathbb{R}^{|K|}$ be a function defined on $(0, 1]$. Player 1 can guarantee $\limsup_{\lambda \rightarrow 0^+} w_\lambda$ in the stochastic game Γ if w_λ satisfies:*

- i) for some integrable function $\phi : (0, 1] \rightarrow \mathbb{R}_+$, $\|w_\lambda - w_{\lambda'}\|_\infty \leq \int_\lambda^{\lambda'} \phi(x) dx$, $\forall 0 < \lambda, \lambda' < 1$;*
- ii) for every $\lambda \in (0, 1)$ sufficiently small, $\Phi(\lambda, w_\lambda) \geq w_\lambda$.*

Remark. In the construction of ε -optimal strategy in Mertens and Neyman [5], w_λ is taken to be v_λ , so condition *i*) is implied by the bounded variation property of v_λ and condition *ii*) is implied by Equation (1.1).

An *absorbing state* is such that once reached, the probability of leaving it is zero. Without loss of generality one assumes that at any absorbing state, the payoff is absorbing (equal to the value of the static game to be played after absorption), as long as one states that both players are informed of the current state.

$\Gamma = \langle K, I, J, g, q \rangle$ is an *absorbing game* if all states but one are absorbing. Mertens, Neyman and Rosenberg [6] used the above Theorem 1.2 to prove the existence of uniform value for absorbing games with finite state space and compact action sets.

Recursive games, introduced by Everett [3], are stochastic games where the stage payoff is always zero before reaching an absorbing state. This note proves the existence of uniform value for a recursive game with finite state space and compact action sets, using an approach analogous to Mertens, Neyman and Rosenberg [6] for absorbing games. Moreover, due to the specific payoff structure, we show that ε -optimal strategies in recursive games can be taken stationary. This is not the case for a general stochastic game, in which an ε -optimal strategy has to be usually a function of the whole past history (cf. Blackwell and Ferguson [2] for the "Big match" as an example).

Everett [3] proved the existence of stationary ε -optimal strategies for both players which guarantee the "limiting-average value" (property (B) in Definition 1.1). As our proof relies

on his characterization of this value (and on its existence), we describe here the result.

Given $S \subseteq \mathbb{R}^d$, let \overline{S} denote its closure.

Let $K^0 \subseteq K$ be the set of nonabsorbing states.

$\Phi(0, \cdot)$ refers to the operator $\Phi(\lambda, \cdot)$ with $\lambda = 0$ in Equation (1.2).

Theorem 1.3 (Everett [3]) *Let Γ be a recursive game with finite state space and compact action sets. Then Γ has a limiting-average value v and both players have stationary ε -optimal strategy, i.e. $\forall \varepsilon > 0$, there are stationary strategies $(\sigma^*, \tau^*) \in \Sigma \times \mathcal{T}$ s.t.: for any $(\sigma, \tau) \in \Sigma \times \mathcal{T}$,*

$$\mathbb{E}_{\sigma^*, \tau}^{k_1} \left[\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n g_t \right] \geq v(k_1) - \varepsilon \quad \text{and} \quad \mathbb{E}_{\sigma, \tau^*}^{k_1} \left[\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n g_t \right] \leq v(k_1) + \varepsilon$$

Moreover, the limiting-average value v is characterized by¹ $\{v\} = \overline{\xi^+} \cap \overline{\xi^-}$, where

$$\begin{aligned} \xi^+ &= \left\{ u \in \mathbb{R}^{|K^0|} : \Phi(0, u) \geq u, \text{ and } \Phi(0, u)(k) > u(k) \text{ whenever } u(k) > 0 \right\}, \\ \xi^- &= \left\{ u \in \mathbb{R}^{|K^0|} : \Phi(0, u) \leq u, \text{ and } \Phi(0, u)(k) < u(k) \text{ whenever } u(k) < 0 \right\}. \end{aligned}$$

Everett [3]'s proof of the above result consists of the following two arguments: first, any vector $u \in \xi^+$ (resp. $u \in \xi^-$) can be guaranteed by player 1 (resp. player 2); second, the intersection of ξ^+ and ξ^- is nonempty.

2 Main results and the proof

We prove that v (as characterized in Theorem 1.3) is also the uniform value of Γ , and that players have stationary ε -optimal strategies.

Theorem 2.1 *A recursive game with finite state space and compact actions sets has a uniform value. Moreover, both players can guarantee the uniform value in stationary strategies.*

Remark. We emphasize that our definition of uniform value includes that of limiting-average value, thus our results extend Everett [3] to a much stronger set-up.

Proof. We first prove that v is the uniform value of Γ using Theorem 1.2. Let u be any vector in ξ^+ . An equivalent characterization for u is:

$$\Phi(0, u) \geq u \text{ and } u(k) \leq 0 \text{ whenever } \Phi(0, u)(k) = u(k), \forall k \in K^0.$$

Define $w_\lambda = u$, $\forall \lambda \in (0, 1)$. We check that w_λ satisfies the two conditions *i*) & *ii*) in Theorem 1.2.

i) It is trivial since u does not depend on λ ;

ii) The crucial point is that $\Phi(\lambda, u) = (1 - \lambda)\Phi(0, u)$ for recursive games. Then the condition $\Phi(\lambda, w_\lambda) \geq w_\lambda$ for all λ close to 0 is satisfied in the following two cases:

- for $k \in K^0$ with $\Phi(0, u)(k) > u(k)$, we have $(1 - \lambda)\Phi(0, u)(k) > u(k)$ for λ close to 0⁺;
- for $k \in K^0$ with $\Phi(0, u)(k) = u(k)$, we have $u(k) \leq 0$, thus

$$(1 - \lambda)\Phi(0, u)(k) = (1 - \lambda)u(k) \geq u(k), \quad \forall \lambda \in (0, 1).$$

¹When working with the operator $\Phi(0, \cdot)$ or $\Phi(\lambda, \cdot)$, it is sufficient to consider those vectors $u \in \mathbb{R}^K$ identical to the absorbing payoffs on the absorbing states $K \setminus K_0$. Whenever no confusion is caused, we identify u with its projection on $\mathbb{R}^{|K^0|}$.

Now Theorem 1.2 states that player 1 can uniformly guarantee any $u \in \xi^+$. A symmetric argument proves that player 2 can uniformly guarantee any vector $u \in \xi^-$. As $\{v\} = \overline{\xi^+} \cap \overline{\xi^-}$ is nonempty by Theorem 1.3, this proves that v is the uniform value of Γ .

Next, we point out that the ε -optimal strategy appearing in the Mertens and Neyman theorem [5] can be taken *stationary* when Γ is a recursive game. Indeed, let w_λ be the function satisfying conditions *i*) & *ii*) in Theorem 1.2. The general construction of σ for player 1 to guarantee $\limsup_{\lambda \rightarrow 0^+} w_\lambda$ is as follows: at each stage $t \geq 1$,

- step 1. the t -stage history is used to compute in an inductive way a small enough discount factor λ_t ;
- step 2. play some optimal strategy $x_{\lambda_t}(w_{\lambda_t}, k_t) \in \Delta(I)$ in the zero-sum game $\Phi(\lambda_t, w_{\lambda_t})(k_t)$, which is defined in Equation (1.2) by letting $f = w_{\lambda_t}$, $\lambda = \lambda_t$, $k = k_t$.

Since 1) we have chosen in the first part of the proof $w_\lambda = u$ to be constant, 2) $\Phi(\lambda, u) = (1 - \lambda)\Phi(0, u)$ for recursive games, this implies that $x_{\lambda_t}(w_{\lambda_t}, k_t)$ can be taken to be some $x(u, k_t)$, an optimal strategy of player 1 in the zero-sum game $\Phi(0, u)(k_t) = \text{val}_{x \in \Delta(I), y \in \Delta(J)} \mathbb{E}_{x, y}^{k_t}[u(k')]$. Hence, to define σ , we do not need (as in Mertens and Neyman [5]) the whole t -stage history to compute λ_t , and the only necessary information is the current state k_t . ■

3 Concluding remarks

1. To have a better understanding of the existence of stationary ε -optimal strategy in recursive games, one can compare our construction to the one in Mertens, Neyman and Rosenberg [6] for absorbing games. Indeed, they have also chosen w_λ to be some constant function u . However, there is no such equality $\Phi(\lambda, u) = (1 - \lambda)\Phi(0, u)$ for absorbing games, so the optimal strategy $x_{\lambda_t}(w_{\lambda_t}, k_t)$ at each stage t for $\Phi(\lambda_t, u)$ depends on λ_t , hence on the whole history.

2. Since the constructed ε -optimal strategy is stationary, our proof extends to *recursive games with signals on actions*: there is no need to assume perfect observation of the opponent's actions, as long as both players are informed of the current state.

3. A weaker notion corresponds to the existence of an *asymptotic value*, i.e. $\lim_{n \rightarrow \infty} v_n = \lim_{\lambda \rightarrow 0} v_\lambda = v$ and was studied in previous works. For absorbing games with finite state space and compact action sets, Rosenberg and Sorin [8] characterized and proved such existence, via the so-called *operator approach*. Using the same approach, Sorin [?] provided the corresponding result for recursive games with finite state space and compact action sets and in particular, showed that the asymptotic value is equal to the unique vector characterized by $\overline{\xi^+} \cap \overline{\xi^-}$ in Everett [3].

However, these convergence results do not extend to general stochastic games with finite state space and compact action sets: Vigeral [11] provided an example with no convergence of v_λ . This implies a fortiori that the existence result of a uniform value for absorbing/recursive games with compact actions does not extend to general stochastic games.

4. Our proof does not extend to *recursive games with infinite state space*. In fact that we use the nonemptiness of $\overline{\xi^+} \cap \overline{\xi^-}$, which is obtained in Everett [3] by an inductive proof on the (finite) number of states.

Li and Venel [4] provided a sufficient condition for recursive games with infinite state space to have a uniform value, that is, the family of n -stage values $\{v_n\}$ being totally bounded for the uniform norm. They presented also an example to have no uniform value when this condition is not satisfied.

Acknowledgement The authors are very grateful to Guillaume Vigeral for helpful comments. Part of Xiaoxi Li's work was done when he was an ATER fellow at THEMA of Université Cergy-Pontoise during the academic year 2015-2016.

References

- [1] T. BEWLEY AND E. KOHLBERG, *The asymptotic theory of stochastic games*, Mathematics of Operations Research, 1:197–208, 1976.
- [2] D. BLACKWELL AND T.S. FERGUSON, *The big match*, Annals of Mathematical Statistics, 39:159–163, 1968.
- [3] H. EVERETT, *Recursive games*, in Contributions to the Theory of Games III, M. Dresher, A.W. Tucker and P. Wolfe (eds.), Annals of Mathematical Studies 39: 47–78, Princeton University Press, 1957.
- [4] X. LI AND X. VENEL, *Recursive games: uniform value, tauberian theorem and the Mertens conjecture "Maxmin = $\lim_{n \rightarrow \infty} v_n = \lim_{\lambda \rightarrow 0} v_\lambda$ "*, International Journal of Game Theory, 45:155–189, 2016.
- [5] J.F. MERTENS AND A. NEYMAN, *Stochastic games*, International Journal of Game Theory, 10:53–66, 1981.
- [6] J.F. MERTENS, A. NEYMAN AND D. ROSENBERG, *Absorbing games with compact action spaces*, Mathematics of Operations Research, 34:257–262, 2009.
- [7] J.F. MERTENS, S. SORIN AND S. ZAMIR, *Repeated games*, Cambridge University Press, 2015.
- [8] D. ROSENBERG AND S. SORIN, *An operator approach to zero-sum repeated games*, Israel Journal of Mathematics, 121:221–246, 2001.
- [9] L.S. SHAPLEY, *Stochastic games*, Proc. Nat. Acad. Sci., 39:1095–1100, 1953.
- [10] S. SORIN, *The operator approach to zero-sum stochastic games*, Chapter 27 in Stochastic Games and Applications, A. Neyman and S. Sorin (eds.), pages 417–426, Kluwer Academic Publishers, 2003.
- [11] G. VIGERAL, *A zero-sum stochastic game with compact action sets and no asymptotic value*, Dynamic Games and Applications, 3:172–186, 2013.